THE ROLE OF EARLY LIFE SOCIOECONOMIC STATUS IN FEMALE BREAST CANCER INCIDENCE

by

Jared Thomas Butler

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Geography

The University of Utah

December 2013

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Jared Thomas Butler**

has been approved by the following supervisory committee members:

Kevin A. Henry	, Chair	**07/01/2013** Date Approved
Simon Brewer	, Member	**07/01/2013** Date Approved
Antoinette M. Stroup	, Member	**07/01/2013** Date Approved

and by **Andrea R. Brunelle** , Chair/Dean of

the Department/College/School of **Geography**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

The effect of affluence upon increased breast cancer incidence is a widely studied phenomenon. However, utilizing affluence at time of diagnosis may prove non-informative due to residential mobility and variable socioeconomic status (SES) through an individual's life. Examining a cohort of women born in Utah from 1945-1959, this study seeks to determine whether individual and/or area-based SES at birth is associated with female breast cancer risk in life, and to determine if the incidence of female breast cancer is spatially clustered based on residential location at birth.

We utilized Cox proportional hazards (PH) models as a means of determining the impact of SES at birth and adult breast cancer incidence. To examine the potential for spatial clustering patterns at birth, space time scan statistics were run employing 1960 census tracts based on cohort members' residence at birth.

Cox PH modeling found that women born into low SES families were less likely to develop breast cancer than women born into the highest SES groups (Q1 HR=0.83 95% CI: 0.72-0.97; Q2 HR=0.81 95% CI: 0.69-0.96). Spatial clustering was limited, though women born in a 2-year time period in South Salt Lake City and the Sugarhouse neighborhood of Salt Lake City exhibited significantly higher risk than their peers (*RR*=2.2, *p*-value 0.005).

This study seeks to identify mechanisms linking SES at birth to adult breast cancer onset. The findings stress the importance of living and economic conditions at birth and their long-term impacts on breast cancer risk factors and incidence.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION AND LITERATURE REVIEW

1.1 Breast Cancer

Breast cancer is one of the most common cancers among American women with approximately 225,000 new cases diagnosed annually (ACS, 2011). Breast cancer is also the second leading cause of cancer death among women in the U.S. (only behind lung cancer); therefore, understanding risk factors for breast cancer is important for targeted interventions.

Studies examining female breast cancer have yielded a number of possible risk factors. Biological factors such as genetic disposition, young age of menarche, and breast density are known factors that increase breast cancer risk (Jeffreys, Warren, Grunnell, McCarron, & Smith, 2004). Other social and economic risk factors including high levels of education and income have also been established (Clarke et al., 2002) as high socioeconomic status (SES) is linked to low parity (e.g., fewer children) and delayed child birth, which increases breast cancer risk (ACS, 2011). In addition, high SES is correlated with high mammography screening utilization, thus increasing early detection and potentially increasing incidence rates (Zapka, Stoddard, & Costanza, 1989).

The effects of affluence upon breast cancer incidence are described in studies examining the high rates of breast cancer in Marin County, California (Clarke et al., 2002; Jacquez et al., 2011). In Marin County, a wealthy suburban county in the San Francisco Bay Area, women experienced a 37% increase in breast cancer incidence over the past decade, substantially higher than the 3% rise found in neighboring regions (Clarke et al.). While no environmental exposures have been identified in Marin County, the SES of the residents in this region is substantially higher than the national average. Furthermore, the majority of migrants into the county were primarily from other affluent regions such as Long Island in New York (Jacquez et al.). Compounding the influx of wealthy women into Marin County is the host of factors that made this population as a whole more susceptible to having higher than expected rates of breast cancer. For example, a large number of women aged 45-64, who may have had children at younger ages, migrated out of the county over the decade of study. This left a large proportion of the cohort as younger, perhaps professional women without children or who delayed child birth, thus resulting in a higher risk of breast cancer. Support for this phenomenon can be seen through the difference between Marin County and California in their proportions of live births to women aged 30 years or older as compared with those aged younger than 30 years (1.60). Therefore, it is likely that the high risk of breast cancer among women in Marin County was not due to some unknown environmental cause, but instead was due to a subpopulation of women with a high prevalence of risk factors for breast cancer (Clarke et al.).

1.2 Spatial Epidemiology and Disease Mapping

The study of geographic variations in disease and health has long provided clues about demographic, environmental, behavioral, socioeconomic, genetic, and infectious risk factors that may affect disease risk and health outcomes. The first examples from the disciplines of epidemiology and geography date back to the 1800s when the locations of affected individuals were plotted on maps to characterize the spread and possible causes of infectious diseases like cholera (Lawson & Williams, 2001). The first examples of disease mapping of noninfectious diseases in relation to environmental factors did not appear until the end of the 19th century with the mapping of cancer incidence in Europe (Lawson & Williams). These include Haviland's 1870 map of cancer rates in Britain, Power's map of cancer cases in a small British village, and Green's 1908 map of cancer incidence in relation to air pollution from coal in France (Koch, 2011). Today, disease maps combined with modern statistical and spatial analytic techniques are routinely used in epidemiologic investigations to describe geographic variations of disease, to generate hypotheses about etiology, and to inform public health practice and resource allocation. There also exists many new and unique opportunities to investigate geographic variations in disease due to advancements in these methods, the development of GIS technologies, and readily available geographically referenced population-based disease data.

While modern techniques utilized in spatial epidemiology and medical geography allow for more efficient and sophisticated methods, several limitations have emerged. For example, data used for population-based spatial epidemiological studies or disease mapping are often derived using information collected and/or measured at the time a disease is diagnosed. This is especially true for large population-based disease surveillance datasets like cancer data that only collect residential addresses at the time of diagnosis. These geographic data are often used in studies as a proxy for environmental and/or socioeconomic risk factors (Galobardes, Lynch, & Smith, 2007), which fails to account for potential exposures that occurred prior to one’s residence at cancer diagnosis and lead to exposure misclassification and attenuation bias (Boscoe, 2011; Jones, Shih, Thurston, Ware, & Cole, 1993; Schulman, Selvin, Shaw, & Malcoe, 1993). This data limitation is especially relevant for analysis of cancers that have long latency periods between exposure and diagnosis. In cases where there are high levels of residential mobility (which leads to greater error in exposure measurement), geographic information at the time of diagnosis is rendered almost noninformative (Manjourides & Pagano, 2011).

The idea that residential histories are important factors to consider when examining cancer risk is not new. In fact, there are many examples of case-control studies of cancer that specifically collect residential histories to ascertain prior carcinogenic exposures across the life span of an individual (Cantor et al., 1987; Darby et al., 2005; Jeffreys et al., 2004; Meliker et al., 2010). Others have collected residential history to ascertain prior socioeconomic risk factors that could have led to increased cancer risk or improved surveillance and screening behaviors (Pudrovska & Anikputa, 2012). Case control studies, however, require extensive interviews to collect information, which increase cost and have had limited success in finding associations between cancer occurrence and residential locations due to recall bias, which is common among studies that collect patient-reported retrospective information, small sample sizes, and inadequate epidemiologic and geographic statistical methods (e.g., disease mapping) to appropriately account for covariates that vary in space and time (Boscoe, 2011).

Despite these challenges, SES has been shown to be associated with health outcomes. Measuring SES in health research not only allows for the description and monitoring of social distributions of a disease, but it also helps to generate hypotheses related to the possible mechanisms through which SES generates health differences

(Galobardes et al., 2007). Generally, lower SES is associated with poorer health, more co-morbidity, and higher mortality compared to higher SES. However, some diseases such as breast cancer and melanoma have been shown to be correlated with affluence (Singh, Miller, Hankey, & Edwards, 2003).

The literature surrounding SES's link to cancer incidence varies markedly in the unit of analysis, as some studies evaluate SES based on individual-level measures and other use area-based summaries of SES. Individual level socioeconomic data are collected through interviews or are obtained from historical or current administrative records. In large population-based health datasets, individual level SES is often not available and is substituted with an area-based measure. Area-based measures focus on social and economic conditions (e.g., percent of persons living below poverty) that affect all individuals who share a social or specific geographic environment (Shavers, 2007). While some researchers use area-level SES measures as a proxy for individual measures, others consider an area-based measure as a unique construct with its own characteristics that can influence health of disease. When area-level measures are used as proxies for individual-level SES, the association of SES and health is often underestimated due to measurement errors arising from giving all individuals in an area the same values (Galobardes et al., 2007). Commonly used SES measures include level of education, occupation, unemployment, income and wealth, and housing type. The use of residential histories to approximate locations for area-based socioeconomic risk factors for cancer, however, requires a robust theoretical framework from which to examine their relationship – one that exploits the very nature of longitudinally-based residential history data and changes in underlying exposures over time. For this study, I propose to use the life course approach to examine the spatial epidemiology of breast cancer.

1.3 Life Course Approach in Breast Cancer Epidemiology

The life course approach posits that the timing of events over the life span is an important determinant of disease risk and outcomes and proposes that exposures at different stages and/or their accumulation over time affect the etiology of disease (Ben-Shlomo & Kuh, 2002; Pickles, Maughan, & Wadsworth, 2007). There are two main life course causal models: risk accumulation and critical period. The critical period model posits that an exposure that occurs at a specific point in time has a direct and lasting effect upon on the physical structure of organs or their functions that are not modified in any dramatic way by later experience (Ben-Shlomo & Kuh). This model represents the idea that events that occur at some critical point in time, regardless of later improvements or decreases in health, directly impact the likelihood of an event at later period of time. In contrast, the accumulation of risk model proposes that some factors cumulatively raise disease risk over the life course. These may include environmental or behavioral insults that gradually damage health in either independent ways, or may cluster together in socially patterned ways (Ben-Shlomo & Kuh). While these two approaches vary in their influence towards future health outcomes, assuming a life course perspective may enhance studies of late-onset health outcomes such as breast cancer.

Life course epidemiology has contributed numerous studies examining associations between conditions at different life stages and subsequent health (Curtis, Southall, Congdon, & Dodgeon, 2004; Darby et al., 2005; Pudrovska & Anikputa, 2012).

Recently, there has been an increased interest in applying this approach to breast cancer, with particular emphasis on examining critical events in early life that initiate the carcinogenesis of breast cancer (Okasha, McCarron, Gunnell, & Smith, 2003).

The life course approach is most often applied to spatial-temporal positioning vis-à-vis a person's residency, which serves as a proxy for potential environmental factors that can have an influence on disease risk and health outcomes. The relationship between breast cancer and past SES is undeniably complex due to the large number of plausible pathways from exposure to diagnosis. Furthermore, although the exact mechanism in which early life SES affects health in adulthood is largely unknown, recent research has shown that area-based measures of SES in early life is significantly associated with adult health outcomes (Curtis et al., 2004).

Residential histories may also provide data to evaluate spatio-temporal variability in breast cancer incidence. Daikwon Han et al. (2005) utilized residential history of individuals in a case control study to examine variability of cancer clusters over time (1) and attempted to identify clusters of high incidence of breast cancer over the life course. They found clustering of breast cancer based on lifetime residence for cases in relation to controls, and a "substantial degree of spatio-temporal variability in the risk surfaces" (Han et al.).

Utilizing a life course perspective, the examination of a woman's residential history has the potential to bring to light early life risk factors, in particular regarding socioeconomic conditions in early life. For example, Pudrovka and Anikputa (2012) found that higher levels of mother's education and early-life income were associated with greater risk of breast cancer in adulthood. By taking a life course perspective, we can examine whether a woman's initial exposure to affluence or poverty is associated with women's choices and lifestyle characteristics in adulthood. Additionally, examining the changes in affluence in consort with breast cancer incidence may provide important clues about how strong of an impact these early life experiences may have upon health in later life.

2. STUDY AIMS AND HYPOTHESISES

The main objectives of this study are to investigate whether individual-level and area-based SES measured at birth affect a woman's breast cancer risk in adulthood and to investigate the clustering of adult onset cancer cases based on residential location at birth. The specific aims are to:

1. Determine if individual and/or area-based SES at birth is associated with female breast cancer risk later in life; and,
2. Determine if the incidence of female breast cancer is spatially clustered at residential location at birth.

These two aims seek to examine how SES affects breast cancer outcomes from a life course epidemiological approach. The primary study hypotheses for specific aim 1 are:

1.1 Women born in families with high SES have higher risks for developing breast cancer in adulthood compared to women born in lower SES families.

1.2 Women born in more affluent geographic areas have higher risk for developing breast cancer in adulthood compared to women born in less affluent areas.

If the critical period theory is correct, then women born into either higher SES families or areas will have higher risks of breast cancer despite potential changes in SES over the life course. For specific aim 2, the primary study hypothesis is:

2.1 Women born in geographic areas with statistically significantly higher than expected breast cancer incidence might share common risk factors contributing towards increased incidence.

3. METHODOLOGY

3.1 Data and Measures

3.1.1 Study population

The study population was selected from the Utah Population Database (UPDB) to create a large population-based cohort of women born between 1945-1959 in Salt Lake and Weber Counties. UPDB is a research resource with longitudinal data that captures events associated with an individual by linking vital status records (births, deaths, marriage), administrative (drivers licenses), health records (medical billing), and cancer registry data. Initial selection of all children born in Utah from 1945-1969 resulted in 593,691 individuals as potential members of this cohort. Individuals were included in the cohort if they (a) appear as the child on a Utah birth certificate; (b) were born in the years from 1945 to 1969; (c) survived to the age of 18 and resided in Utah; and, (d) are linked to another UPDB record other than their parent's record (i.e., parents' birth certificate) or another UCR cancer record. This last exclusion criterion avoids including children in Utah who lived the rest of their lives in another state. With these selection criteria, the cohort available for this study is based on approximately 441,832 birth records covering the entire state of Utah. Following this selection, it was discovered that birth records from 1960 to 1969 had considerably less data on industry and occupation in comparison to earlier records from 1945-1959. The analyses were therefore restricted to those individuals born from 1945-1959, resulting in 192,464 cohort members. This selection also had the effect of limiting the range of ages in cohort members as higher breast cancer incidence is significantly correlated with increasing age. Additionally, it was discovered that state-wide census tract coverage was not implemented until the 1970 Decennial Census and was only available for Salt Lake and Weber counties for the 1960 Census records. Thus, the analysis was restricted to these two counties in order to obtain both individual and area-based levels of SES for the members of the birth cohort. Although we limited the cohort to only two counties, Utah's population has historically been centered on the Wasatch Front, and in 1960, Salt Lake and Weber counties accounted for approximately 55% of the statewide population, 107,153 birth records from Salt Lake and Weber counties. Finally, this cohort was restricted to Whites due to unavailable Hispanic origin data. The final cohort contained 59,610 members consisting of 39,834 female births in Salt Lake County, 10,905 female births in Weber County, and 8,872 that were unknown. The proposed cohort has been linked to the Utah Cancer Registry, a state-wide cancer registry that began collecting cancer incidence data in 1966. Among the 59,610 females in the cohort, there are 1,735 breast cancer cases.

3.1.2 Residential locations at birth

To capture area-based SES at birth, we utilized geocoded data of the mother's residence at the time of the child's birth. Preliminary analysis showed that approximately 99% of the birth records from 1945-1959 in Salt Lake and Weber Counties had a text-based input for the mother's residence at the time of child's birth. Of these, approximately 85% of the UPDB records had previously been geocoded based on full street address, but completeness rates varied and additional geocoding and location verification was completed for the records in our cohort dataset. For both the regression and spatial analysis, we utilized the location of the mother's residence as a means to assign residential location at time of birth.

3.1.3 Geocoding residential locations

Geocoding full street addresses in our cohort dataset was completed using Tele Atlas street reference files and an ArcGIS 10.0 geocoding engine (US Streets 10.0). An x and y coordinate (Universal Transverse Mercator) for each location were obtained and used for linking the data to census tracts for assignment. While the data were matched to modern-day road networks, this did not produced notable errors as cities in Utah are most often lain out in a grid pattern utilizing a sequential street number systems.

Initially, the datasets were split into Salt Lake and Weber counties due to non-contiguous borders as well as potential confounding addresses due to street naming conventions. Automated geocoding through ArcGIS was conducted on each county and allowed for just over 30% of nongeocoded addresses to be matched. These were then hand reviewed for rough placement errors based on city or county locations. Following this initial run, the remaining data were separated and cleaned to remove specific address types that would not be possible to geocode. These involved such text delimiters as “General Delivery,” “Rural,” “BX,” “Route .. Box,” “R..,” “%..,” and RFD. The geocodable data were then hand edited to correct for spelling errors or erroneous city/county assignment. Additionally, a number of outlying cities and ghost towns were set aside and directly assigned to the census tract in which they were located. This was possible due to the large geographic areas of peripheral census tracts as well as geographic features that would restrict municipal boundaries such as a natural canyon. This included addresses located in the towns of Alta, Bingham, Bingham Canyon, Brighton, Copperfield, Copperton, Garfield, Herriman, Magna, Washington Terrace, and Arsenal Villa. Finally, for the remaining Salt Lake County addresses, cities were renamed to their largest neighboring city if initial geocoding failed to match the address to a point. This method proved viable due to adjacent locations often being incorporated into the larger neighbor at a later date as well as the continuous street grids crossing municipality boundaries. City identifiers for addresses that were relabeled can be found in Table 1.

3.1.4 Area-based measures of SES

All geocoded residential locations were spatially joined to the corresponding decennial U.S. census tract data to obtain area-based measures of deprivation (e.g., poverty, income). Although our cohort of births are from the 1945-1959, we learned that census tracts (CT) were not assigned in Utah until the 1960 Census and that 1960 census tracts were also limited to Salt Lake and Weber counties. Thus, the utilization of this 1960 data potentially introduced some measurement error when assigning 1960 SES values to individuals with births from 1945-1959. Additionally, a measure of poverty was not introduced until the 1970 census with only household income based on income ranges being available. Thus a value of “average household income” was created by taking the number of households within each income range and multiplying it by the midpoint of the income range. Once completed for each group, the values for each range were summed and then divided by the total number of households in the census tract to create a value of average household income. This process was computed for all census tracts in Salt Lake

(88) and Weber (31) counties by the Utah Cancer Registry. The Appendix contains an example table showing the above methodology applied to a Salt Lake County census tract.

3.1.5 Individual measures of SES

Individual-level SES measures were derived from data concerning occupation and industry from birth certificates. The coding of occupation and industry (O/I) relies on using U.S. Census codes. The Utah Population Database performed O/I coding for most birth certificates, resulting in approximately 99% O/I coding for the 1945-1959 cohort. The O/I codes are used as inputs to construct Nam-Powers scores, which classify occupations into an 1-100 interval scale according to their respective median education and income levels and have been used in several large studies of mortality (Smith, Mineau, & Bean, 2002; Steenland, Halperin, Hu, & Walker, 2003). The UPDB presently holds text entries on occupation and industry recorded on all Utah birth certificates with most having been coded into quantitative scores based on the Nam-Powers Socioeconomic Status measure. Individual-level SES at birth is defined as the father's Nam-Powers score whenever possible. If father's Nam-Powers score is unknown or there is no father listed on the birth certificate, mother's Nam-Powers score was used. These were then utilized as both continuous (1-100) and quartile (1-4) rankings for the statistical analysis.

A Spearman's rank correlation coefficient was calculated in order to understanding the dependence of individual and area-based income upon one another. With a coefficient of 0.24 (*p*-value 0.0001), only a small positive correlation between individual and area-based incomes were found between increasing individual-level income and the calculated census tract mean income. This finding suggests that familial levels of affluence could vary widely from the average household income of the individual's census tract of residence. Thus, a substantial amount of individual socio-economic variability could be found in each census tract, such as higher income families residing in the lowest quartile census tract or vice versa.

3.1.6 Other demographic and tumor information

Additional variables available from UPDB and UCR that were utilized include dates of birth, dates of cancer diagnosis, birth weight, presence of siblings in the cohort, vital status (alive/deceased), date of last follow-up , primary site, histology, and family history of cancer (1^{st}-3^{rd} degree relative diagnosed with cancer).

3.2 Statistical Methods

3.2.1 Modeling breast cancer risk

To assess the effect of SES at birth on adult breast cancer, we utilized multivariate Cox proportional hazards models, which is a standard technique used to examine the risk of an event (in this case, breast cancer diagnosis) among a cohort of individuals (birth cohort 1945-1949). Cohort members who died of a noncancer event, were lost to follow up, or were still alive at the end of the study, December 31, 2009, were right censored. Since all cohort members had to have reached adulthood to be included in the study, age for each participant was measure in numbers of months from age 18 to either breast cancer diagnosis or censoring.

Multivariate Cox regression models were applied to examine the impact that SES at birth had upon breast cancer incidence. Breast cancer incidence later in life was identified through a binary variable stating either incidence (1) or nonincidence (0). The models utilized controlled for birth weight (pounds), birth year, and whether a child had a sibling in the same cohort to account for the fact that SES is likely to change with increasing family sizes and to serve as a proxy for "familial heritability." Models were run independently for individual SES (Nam-Powers) and for area-based SES (CT average household income) in addition to utilizing an interaction model accounting for joint effects of SES. Both individual and area values of SES were classified to quartile ranks as a means for comparison between low and high income groups. Both measures utilized the highest quartile of SES was the reference group for which lower affluence groups were compared. All members of the cohort were analyzed together as geographic placement of an individual was only utilized as a means for obtaining area-based SES. Significance tests were evaluated at the $p < 0.05$ level. SAS 9.3 software was utilized for the Cox regression analysis.

3.2.2 Exploration of geographic clustering

To assess the potential of geographic clustering of births that are later diagnosed with adult breast cancer, we utilized space-time spatial scan techniques. A spatial analysis of the cohort's residences at birth was conducted to examine potential clustering of individuals who are later diagnosed with breast cancer. Space-time spatial scan techniques were utilized to examine both the spatial and temporal distribution of these births in Salt Lake and Weber counties based upon their residential census tract at birth.

Spatial scan techniques are commonly used as a means to detect clusters in point processes utilizing a window of flexible sizes (typically circular in space) to compare incidence points with their underlying population at risk. Once a threshold of either statistical significance or maximum population size is reached, a relative risk is calculated for the cluster based upon expected number of cancer cases within a particular region. This expected value is computed by SatScan aggregating each tract in the cluster's expected breast cancer incidence rate to the cohort's global average. Additionally, each tract is independently assigned a relative risk of breast cancer incidence in comparison to all other census tracts. When expanding beyond a solely spatial analysis to incorporate time, the windows gain a new, temporal dimension that concurrently increases longitudinally through time while still spatially scanning the geographic area for incidence and population values.

As a means of comparing outcomes and associated relative risk of breast cancer, cohort members were classified as either having experienced a breast cancer event (1) or of being right censored (0). Time to event was calculated via person-time for each individual utilizing their birth date and either the date of first breast cancer diagnosis or the date of censoring. Person times were then aggregated to the census tract level by month creating a total person-time value (population at risk) for each "tract-month." Each county was analyzed independently due to the noncontiguous nature of their locations.

Poisson-based probability models were utilized as the number of events is presumed to be occurring at an average rate for each tract as well as independently of previous events. Therefore, both a case file and a population file were required for analysis. For the case file, each tract was represented for each month and contained tracts geographic identifier (SGID) and the total number of those individuals that would experience a breast cancer incidence later in life. The populations file also contained all tracts' SGID's as well as a population field containing the previously calculated person-time value. Finally, X/Y Cartesian coordinates were generated for each 1960 census tract polygon's centroid utilizing the NAD 1983 UTM Zone 12N projections and a feature to point transformation. This shapefile was then imported as the coordinate file to be utilized by SatScan.

Nondefault parameters utilized for the analysis included the utilization of ellipsoid shaped spatial window rather than default circular to allow for increased spatial range of catchment coverage. Both "Time Precision" and "Time Aggregation" were set to "Month" due to the aggregation of person-time and breast cancer incidence to this level. The maximum spatial cluster size and maximum temporal cluster size were limited to 25% of the study population and time period to allow for results to be statistically significant without encompassing either too large of a geographic region or time period. Finally, scans were restricted to only areas with high relative risk.

Following analysis upon each county's complete population, the datasets were stratified into 5-year groupings (Group 1: 1945-1949, Group 2: 1950-1954, and Group 3: 1955-1959) and run independently as a means to account for increasing age for those born in earlier time periods. This breakdown of population into subcohorts, while decreasing the power of the results, allowed for an investigation into whether significant clustering found within a county could be explained by age rather than socioeconomic or environmental exposures.

Finally, mapping of each tract's relative risk as well as the geographic location of clusters was performed for allow for visual representation of findings. All spatial clustering analysis was accomplished utilizing SAS 9.3, SatScan 9.1.1, and ArcMap 10.1.

3.2.3 Descriptive analysis of clusters

Following the exploration of geographic clustering, a descriptive analysis of demographic characteristics of the cases found inside the significant clusters compared to the cases outside the clusters was completed. Variables that were examined included in this analysis were average population (# persons and total person time), average Nam-Powers score, average household income, average age of first breast cancer diagnosis, and median age of first breast cancer diagnosis (calculated for cohort members experiencing a breast cancer event only). Creation of these values was accomplished utilizing SAS 9.3.

Table 1 - Geocoding City Reclassification

Original City	**Redefined City**
Bennion	Salt Lake City
Cottonwood	Salt Lake City
Crescent	Salt Lake City
Draper	Salt Lake City
East Crescent	Salt Lake City
East Millcreek	Millcreek
Granger	West Valley City
Granite	Salt Lake City
Hunter	West Valley City
Kearns	Salt Lake City
Rural	Salt Lake City
South Salt Lake	Salt Lake City
Union	Salt Lake City
Welby	West Jordan

4 RESULTS

4.1 SES and Breast Cancer Risk

Women born into lower SES families (identified via NP-Scores) were more likely to be diagnosed with breast cancer than women born into the highest quartile group, with the lowest quartile having a hazard ratio of 0.83 (95% CI 0.72-0.97) and the second quartile being slightly lower at 0.81 (95% CI 0.69-0.96) (Table 2). While the third quartile was also lower than the reference, it was not statistically significant (HR 0.93, CI 0.81-1.08). Area-based measures of SES were not significant for adult breast cancer risk.

Interaction models were inconclusive and were not detailed in this manuscript. However, a few interesting results were observed. While individual SES did not directly affect breast cancer risk, when including it as an interaction term, individuals born in lowest SES census tracts were significantly less likely to later develop breast cancer compared to the reference group (HR 0.64, p=0.0147). This effect weakened as Np-SES increased (CT SES Q2/ Np-SES Q2 HR 0.78, p=0.049; CT SES Q1/ Np-SES Q3 HR 0.85, p=0.0192). Finally, females born in the third quartile of census tract SES as compared to the highest, if born into low SES families (Np-SES Q1), were at a lower risk of breast cancer despite living in more affluent areas (HR 0.58, p=0.0329).

4.2 Space-Time Clustering of Breast Cancer Incidence

Examining each county individually, the space-time clustering models looked at the distribution of breast cancer cases in comparison to the population at risk (aggregate person-time) throughout the region.

The primary location that exhibited statistically significant clustering of breast cancer cases included 17 census tracts in central Salt Lake County from January 1945 to February 1948. Incorporating a portion of Salt Lake City and South Salt Lake, this region was highly urbanized as of the 1960 Decennial Census with 23.8% (9,493 of 39,833) of the total births occurring in this area. There were 122 observed breast cancer diagnosed compared to 61.70 that were expected. The relative risk based on the ratio of observed to expected breast cancer cases for women born in these census tracts was 2.09 (p=0.00001). Figure 1 displays the geographic location within the county as well as the two other areas of high relative risk (Cluster 2: RR=3.54, p-value=0.103; Cluster 3: RR=2.11, p-value=0.549) for the Salt Lake County dataset from 1945-1959.

In addition to the location of each geographic cluster's census tracts, Figure 1 also displays each tract's relative risk based upon breast cancer incidence of individuals born into each tract calculated against the total person-time. The lower three classifications are of those counties with lower than expected risk of adult breast cancer incidence (RR <1.0). The upper two divisions (RR = 1.01-2.00) encompass tracts where women born experienced an increased risk of incidence later in life.

Relative risks for women in Weber County were not significant. The primary cluster discovered for county incorporated 2 census tracts in central Ogden containing a very high relative risk of 24.9; however, this was based on only 5 cases to an expected

0.2 (*p*-value=0.084). All remaining secondary clusters were not significant with *p*-values above 0.9. A map of this analysis is located in the Appendix.

Stratifying the datasets into 3 subcohorts based upon 5-year cohorts did not yield additional significant clusters. Table 3 describes the results for the sub-cohort spatial scan statistical analysis and mapping of each analysis can be found in the Appendix.

4.3 Descriptive Statistics of Significant Clusters

Table 4 describes the population demographics of the women from the clusters found in Salt Lake County from 1945-1949 compared to the women residing outside of each cluster. Women born inside Cluster 1 were slightly older at the time of breast cancer diagnosis, lived in census tracts with higher average income, and were born into high SES families. Secondary clusters 2 and 3 were not statisitcally significant; however, it can be seen that cluster 2 did not follow this patter with both increased individual and area-based income being found outside of the cluster. Cluster 3 found decreased individual income levels but increased income by tract for those within its bounds. Average age at breast cancer diagnosis was insignificant for both secondary clusters.

Table 2 - Multivariate Cox Proportional Hazards

Multivariate Cox Proportional Hazards					
SES Type			**Quartile 1**		
	Range	**N**	**Events**	**HR**	**95% CI**
Individual SES (NP-Score)	2-37	15,433	417	0.83	0.72-0.97
Area-Based SES (Mean HH Income)	\$2,578.81-\$5,256.10	14,783	437	0.99	0.82-1.2
			Quartile 2		
		N	**Events**	**HR**	**95% CI**
Individual SES (NP-Score)	38-46	11,009	306	0.81	0.69-0.96
Area-Based SES (Mean HH Income)	\$5,264.23-\$6,128.80	13,922	431	1.16	0.96-1.4
			Quartile 3		
		N	**Events**	**HR**	**95% CI**
Individual SES (NP-Score)	47-71	15,192	450	0.93	0.81-1.08
Area-Based SES (Mean HH Income)	\$6,203.42-\$7,659.36	14,124	386	1.03	0.85-1.25
			Quartile 4		
		N	**Events**	**HR**	
Individual SES (NP-Score)	72-99	14,139	445	Ref	
Area-Based SES (Mean HH Income)	\$7,681.34-\$12,711.11	7,758	217	Ref	

(NP – Nam-Powers Score, HH – Household, CI – Confidence Interval)

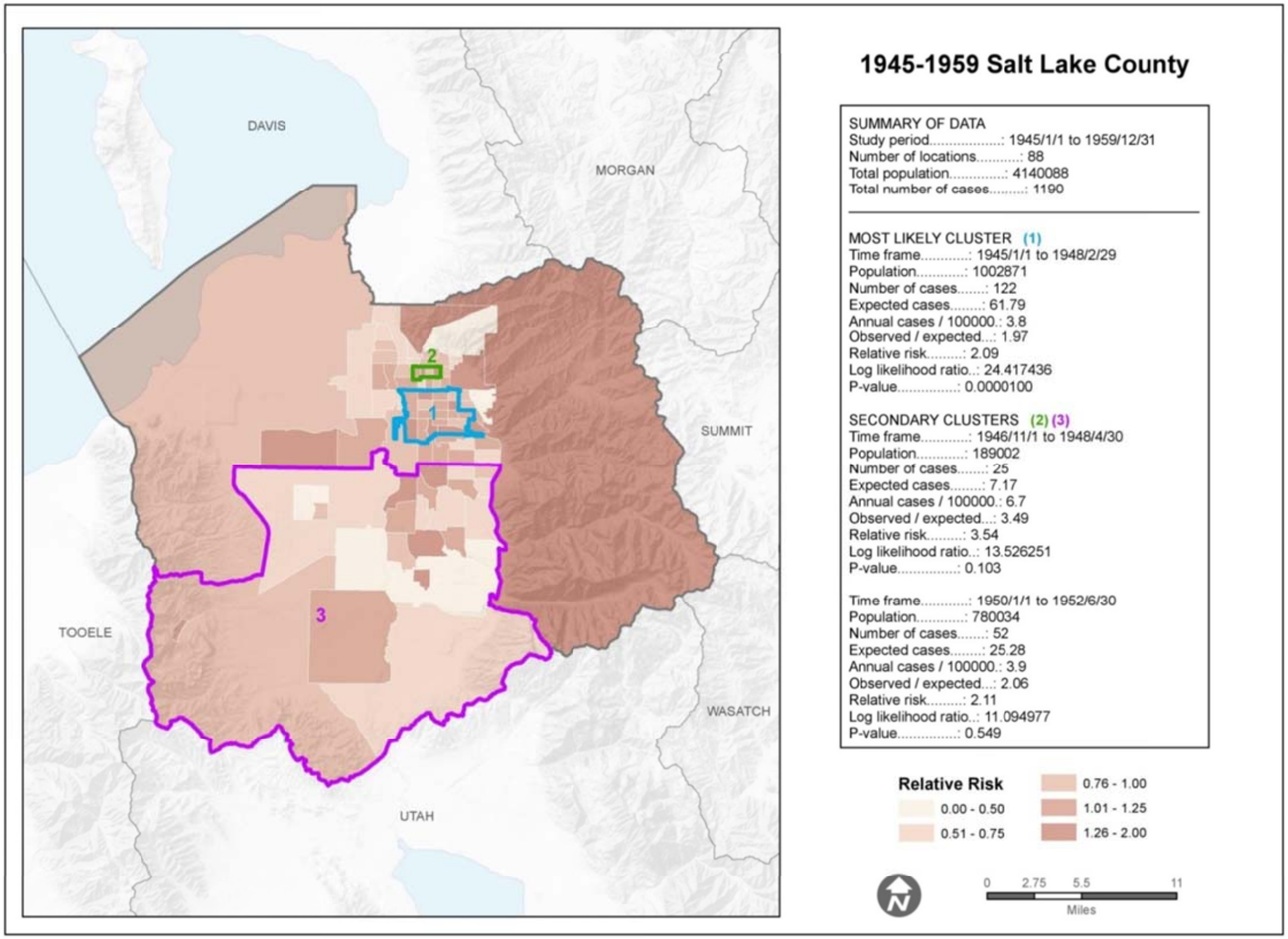

Figure 1 – Map Salt Lake County 1945-1959

Table 3 – Space Time Scan Statistics

Space Time Scan Statistics

Complete Data Sets

	Cluster	# Tracts	Time Start	Time End	Population (Person Time)	% of Total Population	# Cases	Expected Cases	Annual Cases / 100000	Observed / Expected	Relative Risk	Log Likelihood Ratio	P-value
Salt Lake (1945-1949)	1	17	Jan-45	Feb-48	1002871	24.2	122	61.79	3.8	1.97	2.09	24.42	0.00001
	2	4	Nov-46	Apr-48	189002	4.6	25	7.17	6.7	3.49	3.54	13.53	0.103
	3	26	Jan-50	Jun-52	780034	18.8	52	25.28	3.9	2.06	2.11	11.09	0.549
Weber (1945-1959)	1	2	Apr-47	Apr-47	109795	9.7	5	0.2	41.4	24.49	24.9	11.24	0.084
	2	3	Feb-46	Oct-46	178929	15.7	10	2.19	7.7	4.57	4.7	7.5	0.933
	3	8	Mar-49	Aug-49	2602732	22.9	9	1.82	8.3	4.94	5.07	7.29	0.96
	4	6	Nov-47	Nov-47	241435	21.2	4	0.31	22.1	13.11	13.28	6.62	0.997

Stratified Data Sets

	Cluster	# Tracts	Time Start	Time End	Population (Person Time)	% of Total Population	# Cases	Expected Cases	Annual Cases / 100000	Observed / Expected	Relative Risk	Log Likelihood Ratio	P-value
Salt Lake (1945-1949)	1	20	Feb-45	Oct-45	737275	23.2	35	15.22	6.1	2.3	2.42	9.85	0.52
	2	9	Jan-47	Feb-48	499267	15.7	33	15.22	5.7	2.17	2.27	8.15	0.93
	3	5	Nov-46	Aug-47	306785	9.7	23	8.91	6.8	2.58	2.67	7.51	0.93
Salt Lake (1950-1954)	1	2	Sep-54	Sep-54	16463	0.4	3	0.065	96.4	46.2	46.52	8.57	0.89
	2	9	Jul-50	Jan-51	183602	4.4	8	1.26	13.2	6.34	6.44	8.09	0.91
Salt Lake (1955-1959)	1	15	Apr-56	Sep-56	592011	11.7	17	4.22	5.3	4.03	4.19	11.16	0.259
	2	3	May-58	Nov-58	146691	2.9	8	0.99	10.6	8.04	8.22	9.75	0.451
	3	8	Apr-58	Apr-58	598486	11.8	7	0.64	14.4	10.98	11.2	10.47	0.927
Weber (1945-1949)	1	2	Apr-47	Apr-47	99590	11.1	5	0.3	41.4	16.56	17.29	9.44	0.104
	2	4	Jun-49	Jun-49	70081	7.8	3	0.18	40.5	16.23	16.65	5.58	0.952
	3	6	Nov-47	Nov-47	171148	19.1	4	0.45	22.1	8.87	9.16	5.24	0.985
Weber (1950-1954)	1	2	May-52	Jun-52	102982	8.8	4	0.26	25.7	15.66	16.3	7.33	0.67
Weber (1955-1959)	1	2	Jul-56	Sep-56	48252	3.6	3	0.089	40.3	33.71	34.99	7.7	0.81
	2	5	Apr-55	May-56	198592	14.8	10	2.51	4.8	3.99	4.41	6.71	0.82

Table 4 – Descriptive Statistics – Salt Lake County 1945-1959

Salt Lake County Clusters (1945-1959)										
Cluster	N (Persons)	# Tracts	BC Cases	Obs.	Exp.	Relative Risk	Average Nam-Powers Score	Average Income Mid for Tract at Birth	Average Age at Breast Cancer Diagnosis	Median Age at Breast Cancer Diagnosis
Cluster 1	9493	17	342	122	61.79	2.09	55.90*	6137.50*	49.53*	50.33
Outside Cluster 1	30340	71	848				53.26*	6052.00*	48.49*	49.33
Cluster 2	18002	4	73	25	7.17	3.54	49.75*	3897.90*	49.74	51.33
Outside Cluster 2	38031	84	1117				54.07*	6175.40*	48.73	49.42
Cluster 3	32170	26	197	52	25.28	2.11	53.12*	6713.10*	48.35	49.42
Outside Cluster 3	7663	62	993				54.08*	59.19.70*	48.87	49.58

*t-test $p < 0.05$

5 DISCUSSION

5.1 SES Impact on Breast Cancer Incidence

Findings support previous research showing SES early in life has significant effects on the risk of breast cancer in adulthood (Pudrovska & Anikputa, 2012). It was found that individual SES measures obtained from parental occupation and income were stronger indicators of risk when compared to spatially assigned neighborhood SES. Breast cancer risk using SES followed the expected patterns of increasing risk with increasing affluence as seen in previous research (Clarke et al., 2002; Jacquez et al., 2011; Pudrovska & Anikputa). The joint effects of individual and area-based SES on breast cancer risk did indicate that the influences of each are not independent of each other and that area effects were moderated by individual SES. Early life SES and environmental conditions can influence health and breast cancer risk in later life both directly and indirectly through critical period experiences and their influence on future life course trajectories. Direct pathways involve treating the early-life conditions (e.g., low or high SES) as a critical exposure that “directly affects biological processes in childhood or adolescence” (Pudrovska & Anikputa, 2012). This may include such influences as diet or body weight during adolescence, risk factors potentially influenced by parental SES and early life environment. These effects are long lasting and viewed as independent of any considerable changes experience later in life (Ben-Shlomo & Kuh, 2002). In comparison, indirect pathways include the shaping of lifestyle preferences, resulting in potential risk factors such as adult reproductive behaviors, diet, or socioeconomic attainment. Additionally, educational attainment, an important intermediary between SES and breast cancer incidence, can be indirectly shaped through childhood SES due to the potential of a positive association with parental affluence. For example, women born to higher SES families might have a higher risk of breast cancer due to higher educational attainment, a factor commonly associated with increased breast cancer risks such as delays of first childbirth, low parity, and better access to screening services.

Future studies examining breast cancer risk would benefit from data about socioeconomic conditions throughout the life course as well as further information about fertility and healthcare utilization behaviors.

5.2 Space-Time Clustering of Breast Cancer Incidence

Findings from this study suggest possible geographic clustering of the births of women that are later diagnosed with breast cancer in Salt Lake County. Since the primary cluster was found among the oldest women in the cohort (born from 1945-1947), those at greatest risk of breast cancer, an examination of the cohort stratified by 5-year intervals (1945-1949) was also conducted. The census tracts in the cluster were still found by the scan statistic, albeit in two independent regions separated by approximately 2 years. However, the clusters in this region were no longer significant likely due to the smaller number of cases and underlying population. The spatial coherency found between the initial analysis that included the entire cohort (1945-1959) and the analysis stratified into 5-year intervals provide evidence that earlier birth year was not responsible for cluster.

The descriptive statistics of this area compared to Salt Lake County indicate increased levels of parental and area based SES for cohort members in the cluster. However, the higher average age at diagnosis also point towards that this population having been at an increased risk solely due to the earlier births in that region compared to those outside the cluster. Potential influences towards increased SES and therefore increased breast cancer risk may be related to the intra-urban migrations of affluent populations from the city's core to surrounding communities and suburbs. From 1940 to 1950, the population of South Salt Lake City increased from 1,599 to 7,704. The Sugarhouse neighborhood east of South Salt Lake City has traditionally been a highly affluent residential region of the city. Additionally in 1956, Sugarhouse Park was established near the neighborhood's cultural center, providing a 110.5 acre green space for the community's expanding population. This contribution to the community as well as the location's positive reputation may have provided incentive towards residential development and potentially explains the increased average household income and Nam-Powers scores found within the cluster. This grouping of individuals with similar breast cancer risk factors in a particular geographic location is similar to findings within Marin County, California (Clarke et al., 2002; Jacquez et al., 2011).

These findings should be considered with caution as we only found one statistically significant region with an increased breast cancer risk. Additionally, the area detected was a very large section of the populated regions of the county at the time and incorporated a significant population of the cohort. While SatScan did find other areas of potentially higher risk of breast cancer incidence, the low number of cases and underlying population limited the power to detect significant differences.

5.3 Potential Improvements and Further Research

A number of limitations and potential improvements became apparent throughout the research that could be further investigated to improve results found in this or future analysis. Concerning data availability and processing, the assignment of all individuals to 1960 geographic data has the potential to misrepresent the conditions of a tract throughout time. This effect increases as you regress in time with cohort members born earliest having the highest potential measurement error.

Analyses of these findings are for the most part limited to Salt Lake County due to the low levels of population and the resulting lack of significance in Weber county census tracts at this time. While still accounting for a large proportion of the Utah population, at the time, there was only a total population of 890,627. Thus, the number of female births over the course of the study is expectantly smaller than would be found in more densely populated regions of the country around this time.

Due to data being linked to only the Utah Cancer Registry, there exists a strong probability that surveillance bias was a significant issue towards follow up with members of the cohort. As an individual was right censored if they moved from the state at any time prior to the end of follow up, we are most likely missing a number of female births that would later experience a breast cancer event. Linkage of the data set with other state's cancer registries is a potential strategy that would allow for longer follow up and more complete cancer surveillance for the cohort.

While census tracts allowed for substantial investigations into geographic clustering, utilization of actual home locations rather than aggregations could allow for a much finer detail spatial analysis. Rather than the spatial scan window being restricted to pulling from the centroid location of the census tract in which an individual resides, direct residential locations could allow for some members of a tract to be included while others would remain outside of a potential cluster. This is particularly important in more rural census tracts of Utah due to the large geographic areas and low population density commonly found. However, utilization of point position data does introduce a number of confidentiality concerns and would therefore require additional work towards masking identifying information for cohort members.

Further investigation into clustering analysis could be improved through the utilization of additional time subsets for the stratified cohorts. For example, analysis of those cohort members that were born from 1953-1957 (or any other 5-year grouping) could potentially expose a subpopulation of at-risk individuals that were overlooked through those stratifications utilized.

Finally, while this study is restricted to residence of an individual at birth, there is the potential of utilizing a case-control study design as a means to examine the cumulative effects theory of life course epidemiology and how changing SES over time may impact breast cancer incidence. Through collection of multiple residential locations across a woman’s lifetime, both individual and area-based SES trajectories from birth until incidence could allow for greater insight into how changing affluence or neighborhood effects may affect breast cancer risk.

6 CONCLUSION

Few studies have examined the impact of social or environmental risk factors at birth with respect to cancer incidence in adulthood. However, critical events throughout early life are potentially very powerful influences upon future health outcomes as lifestyle choices or direct impacts on health may be formed early on. This study provides further evidence that a family's socioeconomic status at birth contributes to female breast cancer risk. While area-based SES and spatial clustering was not strongly detected in this study, the potential for early area-based influences on breast cancer may exist if populations share similar risk factors. Concurrently, these populations could be identified via spatial clustering analysis if similar influences on breast cancer risk are present for neighborhoods or cities.

Research into the effects of early life on breast cancer incidence later in life should be examined further to gain a better understanding of the impact of critical time periods on health. Utilizing resources such as the Utah Population Database or other long-term information sources has the potential to allow for a life course approach towards breast cancer incidence to expand our understanding on risk factors towards breast cancer development.

APPENDIX

Table 5 – Average Household Income Example Tract

Census Tract: 49003500001			
Income Range	**N Households**	**Mid-Point**	**Total Household Income**
(A)	**(B)**	**(C)**	**(D)=(B*C)**
Less than $1,000	160	$500.00	$80,000.00
$1,000 - $1,999	92	$1,500.00	$138,000.00
$2,000 - $2,999	140	$2,500.00	$350,000.00
$3,000 - $3,999	75	$3,500.00	$262,500.00
$4,000 - $4,999	122	$4,500.00	$549,000.00
$5,000 - $5,999	118	$5,500.00	$649,000.00
$6,000 - $6,999	112	$6,500.00	$728,000.00
$7,000 - $7,999	42	$7,500.00	$315,000.00
$8,000 - $8,999	42	$8,500.00	$357,000.00
$9,000 - $9,999	27	$9,500.00	$256,500.00
$10,000 - $14,999	39	$12,500.00	$487,500.00
$15,000 - $24,999	4	$20,000.00	$80,000.00
$25,000 and over	0	$2,500.00	$0.00
Total Census Tract Households: 973		**Total Census Tract Income: $4,252,500.00**	
Average Household Income ($4,252,500.00/973):			**$4,370.50**

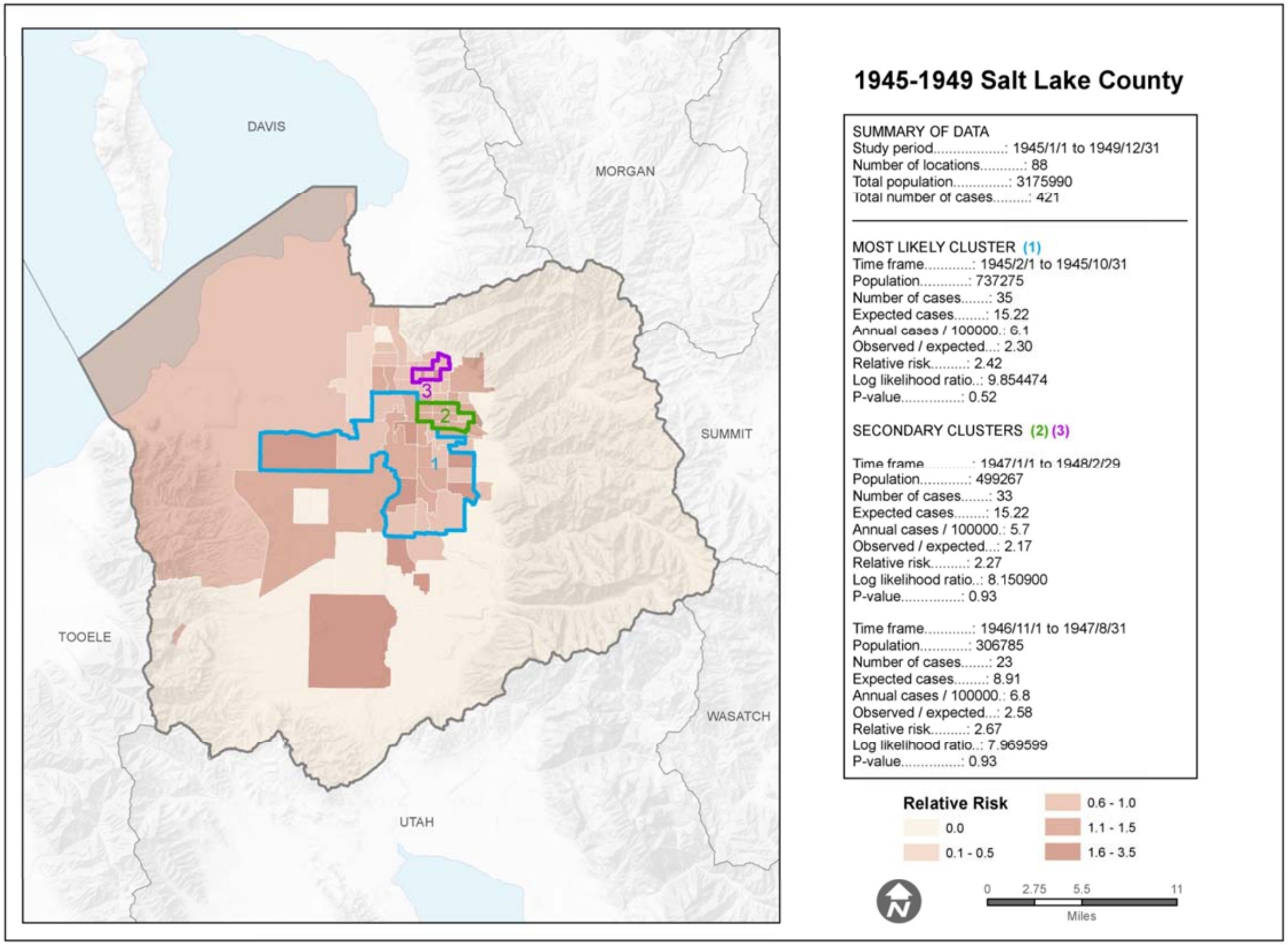

Figure 2 – Map Salt Lake County 1945-1949

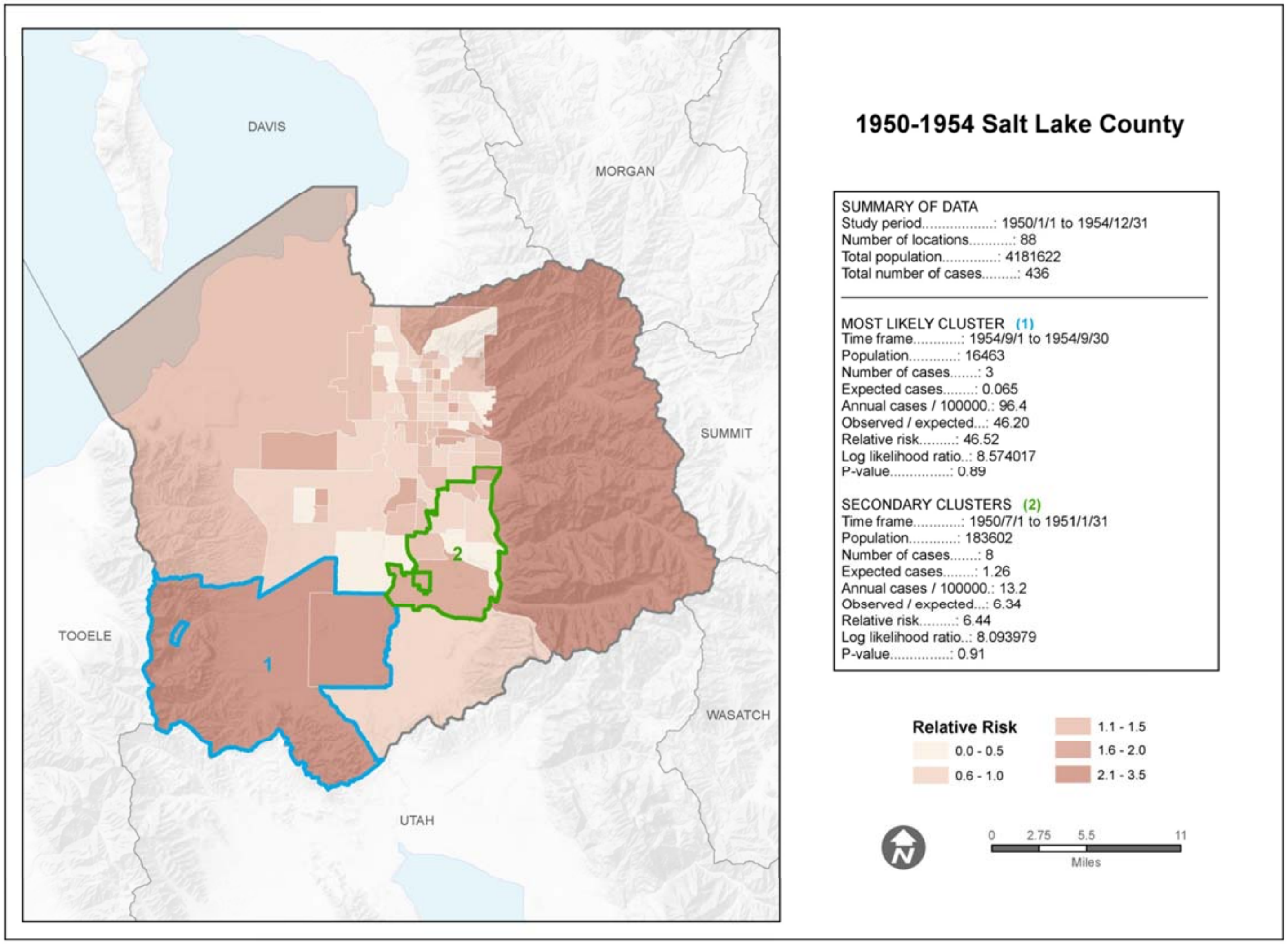

Figure 3 – Map Salt Lake County 1950-1954

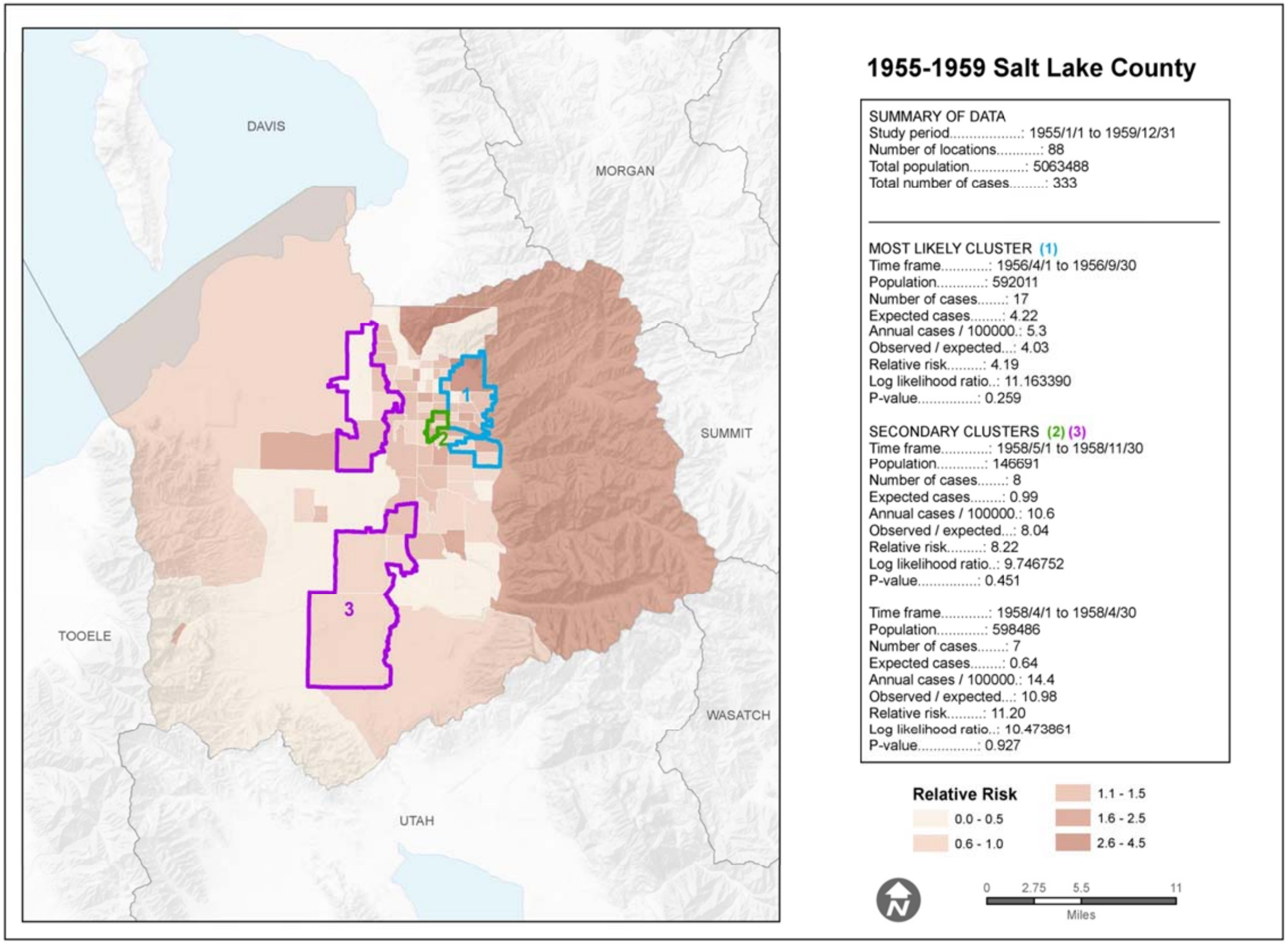

Figure 4 – Map Salt Lake County 1955-1959

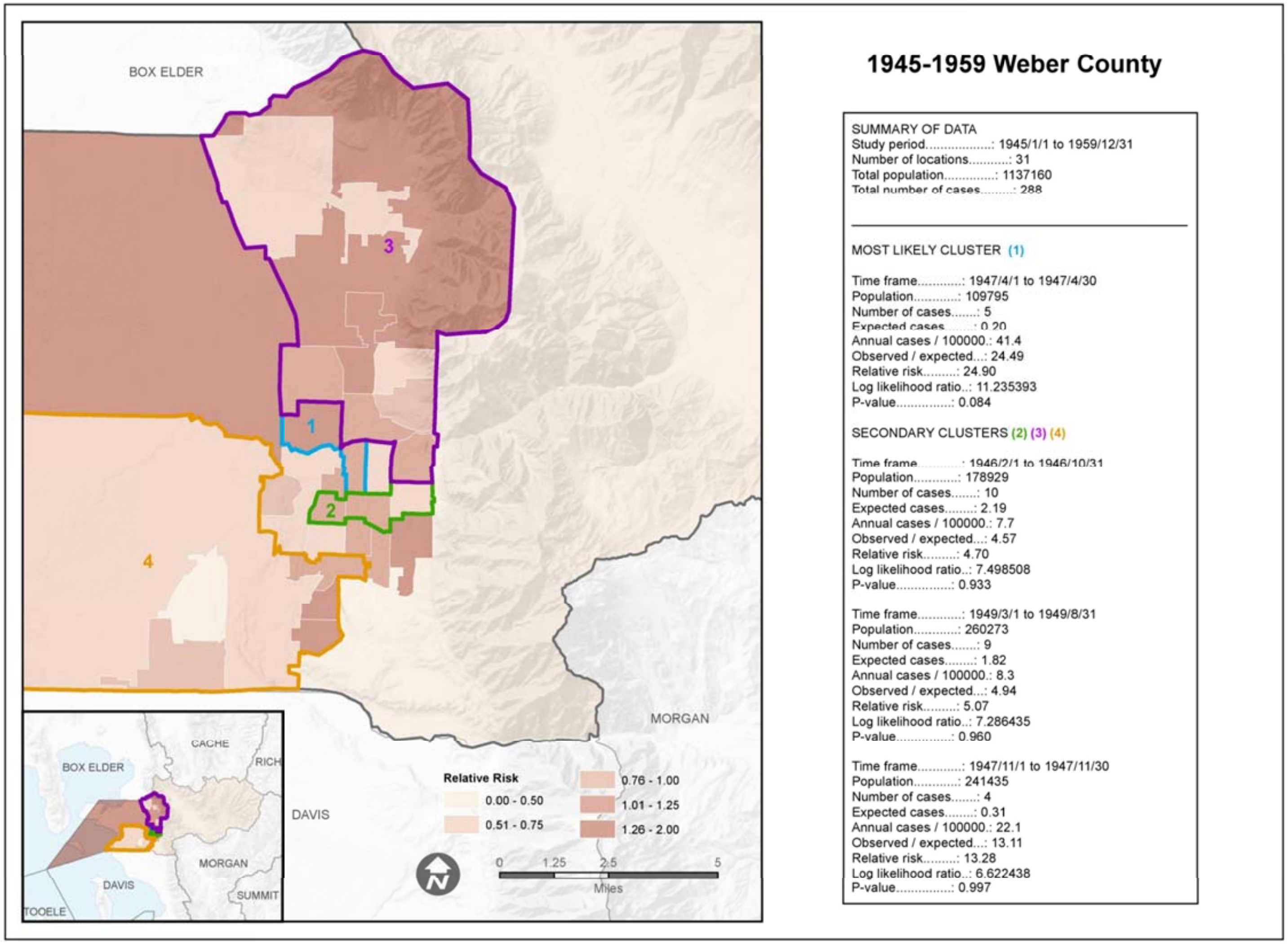

Figure 5 – Map Weber County 1945-1959

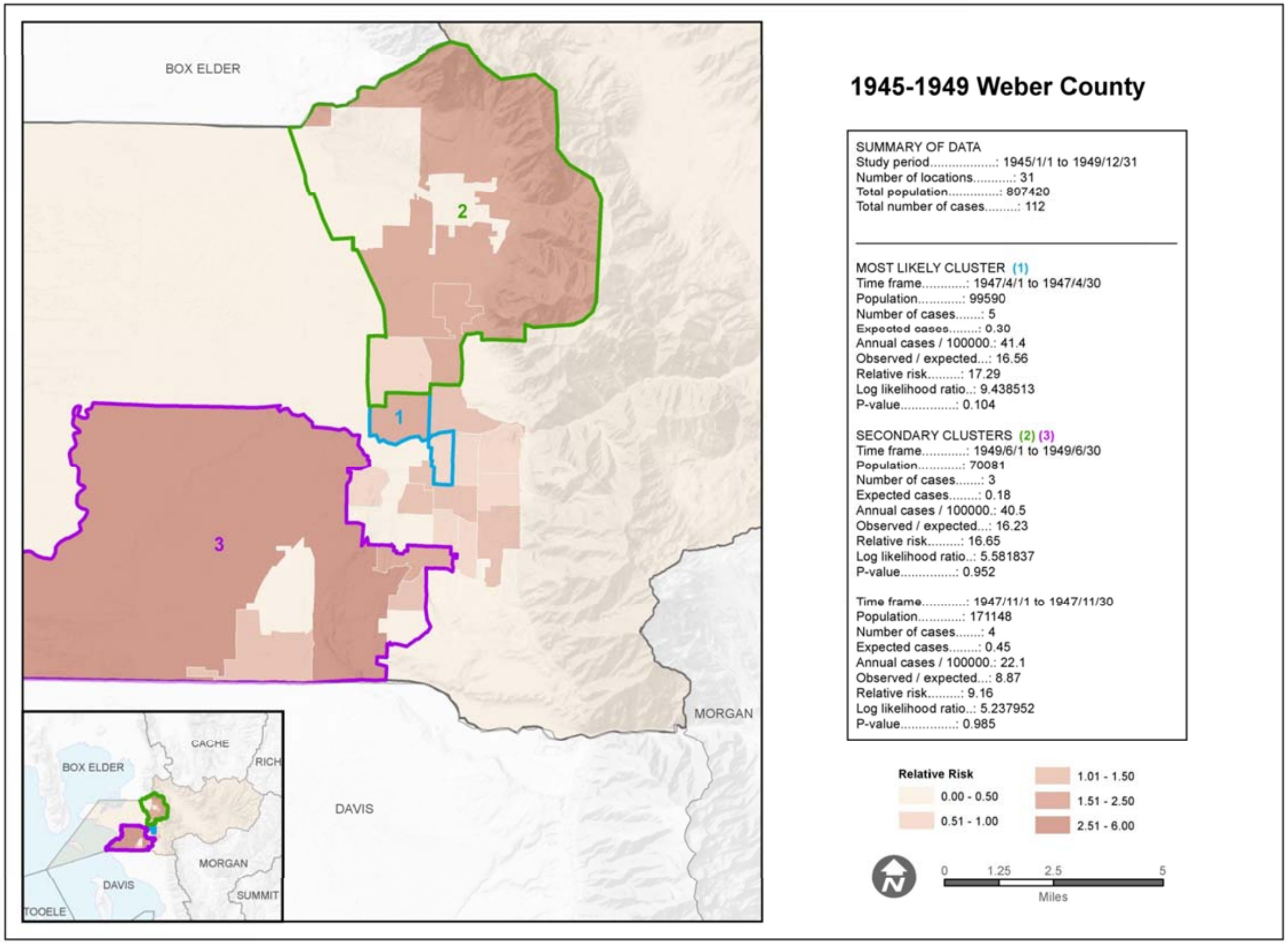

Figure 6 – Map Weber County 1945-1949

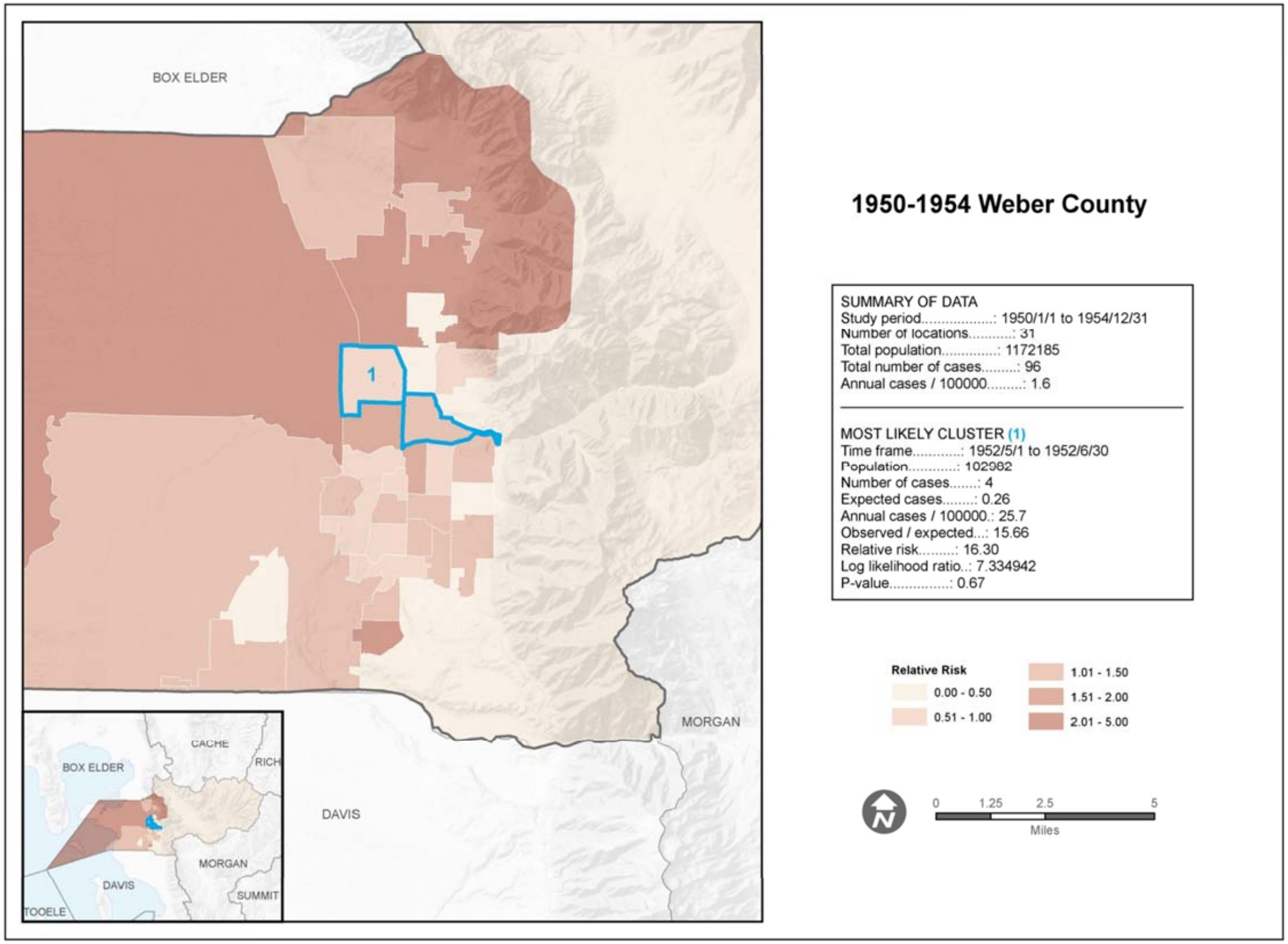

Figure 7 – Map Weber County 1950-1954

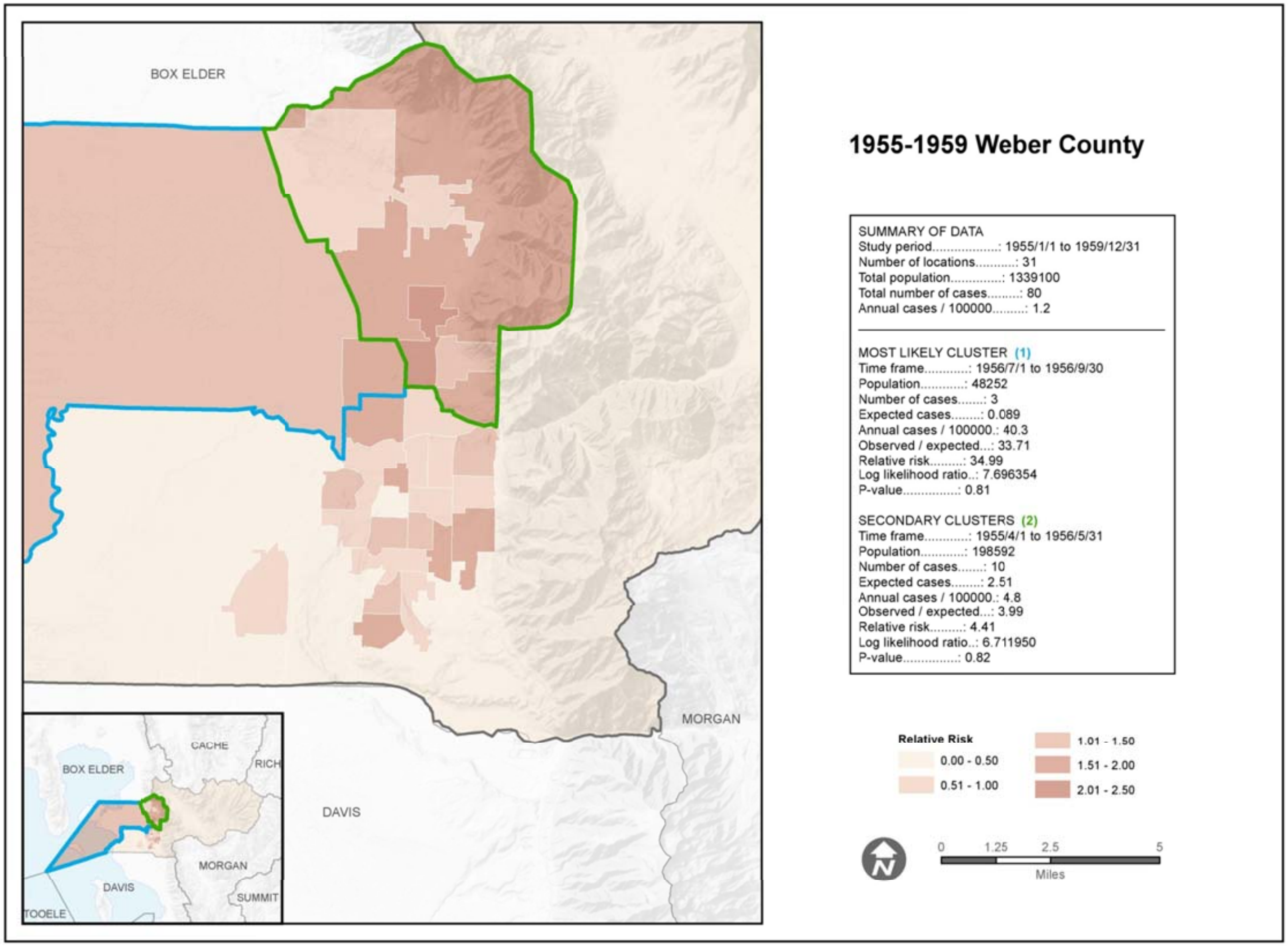

Figure 8 – Map Weber County 1955-1959

REFERENCES

American Cancer Society. (2011). *Breast Cancer Facts & Figures 2011-2012*. Atlanta, GA: American Cancer Society, Inc. http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-030975.pdf

Ben Shlomo, Y., & Kuh, D. (2002). A life course approach to chronic disease epidemiology: Conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology, 31*(2), 285-293. doi: 10.1093/ije/31.2.285.

Boscoe, F. P. (2011). The use of residential history in environmental health studies. In J. Maantay & S. McLafferty (Eds.), *Geospatial analysis of environmental health* (pp. 93-110). New York, NY: Springer.

Cantor, K . P., Hoover, R., Hartge, P., Mason, T. J., Silverman, D. T., Altman, R., . . . West, D. W. (1987). Bladder cancer, drinking water source, and tap water consumption: A case-control study. *Journal of the National Cancer Institute, 79*(6), 1269-79. http://www.ncbi.nlm.nih.gov/pubmed/3480378

Clarke, C. A., Glaser, S. L., West, D. W., Ereman, R. R., Erdmann, C. A., Barlow, J. M., & Wrensch, M. R. (2002). Breast cancer incidence and mortality trends in an affluent population: Marin county, California, USA, 1990-1999. *Breast Cancer Research*, *4*(6), R1. doi: 10.1186/bcr458

Curtis, S., Southall, H., Congdon, P., & Dodgeon, B. (2004). Area effects on health variation over the life-course: Analysis of the longitudinal study sample in England using new data on area of residence in childhood. *Social Science & Medicine, 58*(1). 57-74. doi: 10.1016/S0277-9536(03)00149-7

Darby, S., Hill, D., Auvinen, J. M., Barros-Dios, H., Baysson, H., Bochicchio, F., . . . Doll, R. (2005). Radon in homes and risk of lung cancer: Collaborative analysis of individual data from 13 European case-control studies. *BMJ: British Medical Journal, 330*(7485), 223-226. http://www.jstor.org/stable/25458778

Galobardes, B., Lynch, J., & Smith, G. D. (2007). Measuring socioeconomic position in health research. *British Medical Bulletin, 81-82*(1), 21-37. doi: 10.1093/bmb/ldm001

Han, D., Rogerson, P. A., Bonner, M. R., Nie, J., Vena, J. E., Muti, P., . . . Freudenheim, J. L. (2005). Assessing spatio-temporal variability of risk surfaces using residential history data in a case control study of breast cancer. *International Journal of Health Geographics, 4*(9), 1-9. doi: 10.1186/1476-072X-4-9

Jacquez, G. M., Barlow, J., Kaufmann, A., AvRuskin, G., Rommel, R., & Rasul, J. (2011). Residential mobility and breast cancer in Marin County, California. *International Cartographic Association Workshop on Geospatial Analysis and Modeling.* 1-34. http://www.biomedware.com/?=Publications

Jeffreys, M., Warren, R., Gunnell, D., McCarron, P., & Smith, G. D. (2004). Life course breast cancer risk factors and adult breast density (United Kingdom). *Cancer Causes & Control, 15*(9), 947-955. http://www.jstor.org/stable/3553604

Jones, T. L., Shih, C. H., Thurston, D. H., Ware, B. J., & Cole, P. (1993). Selection bias from differential residential mobility as an explanation for associations of wire codes with childhood cancer. *Journal of Clinical Epidemiology, 46*(6), 545-548. http://www.ncbi.nlm.nih.gov/pubmed/8501481

Koch, T. (2011). *Disease maps: Epidemics on the ground.* Chicago, IL: The University of Chicago Press. doi: 10.1080/14649365.2013.771946

Lawson, A. B., & Williams, F. (2001). *An introductory guide to disease mapping.* Chichester, UK: John Wiley & Sons. doi: 10.1002/0470842571.app

Manjourides, J., & Pagano, M. (2011). Improving the power of chronic disease surveillance by incorporating residential history. *Statistics in Medicine, 30*(18), 2222-2233. doi: 10.1002/sim.4272

Meliker, J. R., Slotnick, M. J., AvRusking, G. A., Schottenfeld, D., Jacquez, G. M., Wilson, M. L., . . . Nriagu, J. O. (2010). Lifetime exposure to arsenic in drinking water and bladder cancer: A population-based case-control study in Michigan, USA. *Cancer Causes & Control, 21*(5), 745-757. doi: 10.1007/s10552-010-9503-z

Okasha, M., McCarron, P., Gunnell, D., & Smith, G. D., (2003). Exposures in childhood, adolescence and early adulthood and breast cancer risk: A systematic review of the literature. *Breast Cancer Research and Treatment, 78*(2), 223-276. doi: 10.1023/A:1022988918755

Pickles, A., Maughan, B., & Wadsworth, M. (2007). *Epidemiological methods in life course research*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780198528487.003.0008

Pudrovska, T., & Anikputa, B., (2012). The role of early-life socioeconomic status in breast cancer incidence and mortality: Unraveling life course mechanisms. *Journal of Aging and Health, 24*(2), 323-344. doi: 10.1177/0898264311422744

Schulman, J., Selvin, S., Shaw, G. M., & Malcoe, L. H. (1993). Exposure misclassification due to residential mobility during pregnancy in epidemiologic investigations of congenital malformation. *Archives of Environmental Health, 48*(2), 114-119. doi: 10.1080/00039896.1993.9938404

Shavers, V. L. (2007). Measurement of socioeconomic status in health disparities research. *Journal of the National Medical Association*, *99*(9), 1013-23. http://www.ncbi.nlm.nih.gov.ezproxy.lib.utah.edu/pmc/articles/PMC2575866/

Singh, G.K., Miller, B.A., Hankey, B.F., & Edwards, B.K. (2003). Area socioeconomic variations in U.S. cancer incidence, mortality, stage, treatment, and survival, 1975-1999. *NCI Cancer Surveillance Monograph Series Number 4* (NIH Publication No. 09-5417). Bethesda, MD: U.S. National Cancer Institute. http://permanent.access.gpo.gov/gpo834/ses_monograph.pdf

Smith, K. R., Mineau, G. P., & Bean, L. L. (2002). Fertility and post-reproductive longevity. *Biodemography and Social Biology, 49*(3-4), 185-205. doi: 10.1080/19485565.2002.9989058

Steenland, K., Halperin, W., Hu, S., & Walker, J. T. (2003). Death due to injuries among employed adults: The effects of socioeconomic class. *Epidemiology, 14*(1), 74-79. http://www.jstor.org/stable/3703282

Zapka, J. G., Stoddard, A. M., Costanza, M. E., & Greene, H. L. (1989). Breast cancer screening by mammography: Utilization and associated factors. *American Journal of Public Health, 79*(11), 1499-1502. http://www.ncbi.nlm.nih.gov/pubmed/ 2817160